

Re: 20TB Storage System

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/performance/2003-09/0015.html>

From: Dan Nelson (dnelson_at_allantgroup.com)

Date: 09/03/03

Date: Tue, 2 Sep 2003 18:29:02 -0500

To: Max Clark <max.clark@media.net>

In the last episode (Sep 02), Max Clark said:

[quoting format manually recovered]

> *Dan Nelson wrote*

> > *Depends on whether you plan on crashing or not :) According to*
> > <http://lists.freebsd.org/pipermail/freebsd-fs/2003-July/000181.html>,
> > *you may not want to create filesystems over 3TB if you want fsck to*
> > *succeed. I don't know if that's using the default newfs settings*
> > *(which would create an insane number of inodes), though.*

>

> *This is a big problem (no pun intended), my smallest requirement is*
> *still 5TB... what would you recommend? The smallest file on the*
> *storage will be 500MB.*

I'd say try formatting a 5TB filesystem with the values you'd use (use a very large `-i`; 1048576 maybe?) and see how much memory fsck consumes. I don't know what UFS2's max blocksize is, but a larger blocksize would help too. You should be able to fake enough storage to do the test with `mdconfig` and some large sparse files.

> > *To sustain only 30MByte/s across the entire set? Doesn't really*
> > *matter, since even a single disk could do that.*

>

> *What would I see better performance with ccd or vinum? So a better*
> *question isn't if I can sustain with 30MByte/s but what would I*
> *expect to maintain?*

For sequential access to mirrored arrays, your bottleneck will probably be the ATA->FC bridges, since they claim to only do 100MBytes/sec. If your three HBAs are 1gbit, then those will be your bottleneck and you'll be able to do 300MB/s reads, and 150MB/s writes (50% mirror penalty). If they're 2gbit and you have 6 bridges, you'll max out at 600MB/s and 300MB/s. If you want to use `vinum raid5`, cut those write speeds in half again (25% `raid-5` penalty).

Theoretically, assuming you can max your FC links and your server can handle the load :) I do `mrtg` graphs of my fibre switches, and I haven't

freebsd-performance: Re: 20TB Storage System

seen it peak over 80MB/sec through a 1gbit link, but I regularly see 70MB/sec sustained to some Tru64 Alpha servers. I only have external hardware raid, though, so I don't know what kind of penalty ccd/vinum will add on top of that. Shouldn't be too much.

--

Dan Nelson
dnelson@allantgroup.com

freebsd-performance@freebsd.org mailing list

<http://lists.freebsd.org/mailman/listinfo/freebsd-performance>

To unsubscribe, send any mail to "freebsd-performance-unsubscribe@freebsd.org"